View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Edinburgh Research Explorer

THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Culture

Citation for published version:
Tamariz, M & Kirby, S 2015, 'Culture: Copying, compression, and conventionality', *Cognitive Science*, vol. 39, no. 1, pp. 171–183. https://doi.org/10.1111/cogs.12144

Digital Object Identifier (DOI):
10.1111/cogs.12144

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Cognitive Science

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 11. May. 2020

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 39 (2015) 171–183
Copyright © 2014 Cognitive Science Society, Inc. All rights reserved.
ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12144

Culture: Copying, Compression, and Conventionality

Mónica Tamariz, Simon Kirby

Language Evolution and Computation, School of Philosophy, Psychology and Language Sciences, The University of Edinburgh

Received 7 June 2013; received in revised form 12 December 2013; accepted 17 December 2013

Abstract

Through cultural transmission, repeated learning by new individuals transforms cultural information, which tends to become increasingly compressible (Kirby, Cornish, & Smith, 2008; Smith, Tamariz, & Kirby, 2013). Existing diffusion chain studies include in their design two processes that could be responsible for this tendency: learning (storing patterns in memory) and reproducing (producing the patterns again). This paper manipulates the presence of learning in a simple iterated drawing design experiment. We find that *learning* seems to be the causal factor behind the increase in compressibility observed in the transmitted information, while *reproducing* is a source of random heritable innovations. Only a theory invoking these two aspects of cultural learning will be able to explain human culture's fundamental balance between stability and innovation.

Keywords: Cultural transmission; Iterated learning; Compression; Imitation; Conventionality

1. Introduction

Social transmission, or social learning, whereby knowledge is passed on generation after generation by imitation, copying, or teaching, is the key process in cultural evolution (e.g., Boyd & Richerson, 1985; Richerson & Boyd, 2005; contributions to Heyes, 2012; Heyes, Huber, Gergely, & Brass, 2009). The role of social transmission as an agent of change in human culture has been widely studied using the transmission or diffusion chain method. In this paradigm, the cultural productions of an individual are used to train the next individual (e.g., Bartlett, 1932; Bangerter, 2000; Esper, 1925; Kalish, Griffiths, & Lewandowsky, 2007; Kashima, 2000; Kirby et al., 2008; Mesoudi, Whiten, & Dunbar, 2006). Repeated observation and production can amplify even weak transmission biases (e.g., a preference for simpler concepts) in such a way that, over time, the transmitted

Correspondence should be sent to Mónica Tamariz, Language Evolution and Computation, School of Philosophy, Psychology and Language Sciences, The University of Edinburgh, 3 Charles Street, EH8 Edinburgh, UK. E-mail: monica@ling.ed.ac.uk

information comes to reflect those biases (e.g., simpler concepts come to prevail; Brighton, Kirby, & Smith, 2005; Kirby, Dowman, & Griffiths, 2007).

We claim that a key adaptation of cultural information in response to social transmission is compressibility. Compressibility is a property of information inversely related to Kolmogorov, or algorithmic, complexity. For a compressible system, it is possible to write a description whose length (in bits of information) is only a fraction of the length of the system itself. The smaller the fraction, the more compressible—and therefore simpler—the system. Compressibility is adaptive for a system that is being culturally transmitted because it makes patterns easier to learn and store in memory (Chater & Vitanyi, 2003) and therefore more likely to persist over cultural time.

A growing number of studies show that cultural systems, when repeatedly learned or used, do become increasingly *simple* or *compressible* (or decreasingly complex); for example, drawings become more schematic (e.g., Galantucci, 2005; Garrod, Fay, Leed, Oberlander, & MacLeod, 2007), and sets of sounds and languages acquire regular structure (Kirby et al., 2008; Perfors & Navarro, 2014; Theisen-White, Kirby, & Oberlander, 2011; Verhoef, Kirby, & Padden, 2011).

Another instance of this adaptation is the fact that repeatedly transmitted knowledge becomes increasingly *conventional*. In his classic theory of remembering, Bartlett (1932) highlighted the influence of "schemata," or existing memory patterns based on experience, on remembering, and reproducing new information. Bartlett used the transmission chain method—which he called "serial reproduction"—in experiments where a participant was exposed to and later asked to reproduce a story or a drawing. One of the conclusions from this and similar experiments is expressed by Bartlett (1932, 185) when he said that the item produced "sooner or later (. . .) tends to assume the form of accepted conventional representations." Bartlett's thinking went to form the basis of cognitive theories of meaning and culture (e.g., Fillmore, 1977; Langacker, 1987; Minsky, 1975; Schank & Abelson, 1977). Sperber's (1996) epidemiology of representations theory also incorporates this bias toward pre-existing, conventional knowledge, which he considered constitute "cultural attractors."

The cultural transmission chain experiments cited above typically involved a participant (generation) being exposed to some data, and then reproducing them. Each participant, therefore, had to do two distinct tasks: learning (holding the information in memory) and reproduction (producing it again). In this study, we manipulate the presence of learning to explore its causal relationship with the observed increase in compressibility (conventionality and simplicity). If the increase in compressibility is caused by repeated reproduction, we expect that our manipulation will have no effect. But if the increase in compressibility is specifically due to holding information in memory, we expect to see an increase in compressibility in that condition only.

2. Methods

We carried out a transmission chain study using a graphical task where chains of participants were exposed to a drawing and immediately were asked to reproduce it as faithfully as possible. In the *Memory* condition, the model drawing is removed from view before reproduction; therefore, the participants were forced to learn it and hold it in memory for a few seconds. In the *Copy* condition, the model drawing remains in full view during reproduction, thus removing the need for learning at all. The requirement for reproduction was similar in both conditions. We then examined how the complexity and conventionality of the drawings changed over time.

2.1. Participants

A total of 308 participants took part in the study (54% female). Given the very short time each participant had to invest in the experiment, we chose to approach students directly during their free time in the courtyards of the faculties of Humanities and Psychology in the University of Seville, Spain and in the Students' Union buildings at the University of Edinburgh. The experimenter introduced herself as a researcher and briefly explained that she was carrying out an academic experiment, unrelated to any sales or advertising purposes, and then asked the potential participant whether he or she would like to participate, stating that the whole thing would take a maximum of 2 min. The few students who showed signs of rejection were dismissed. The rest were invited to sit down in order to proceed.

At the analysis stage, six additional participants were paid £10 each for judging all the drawings.

2.2. Materials

Each transmission chain was initialized with one of the drawings shown in Fig. 1, drawn on a white A5 card with a black *edding 1,200* felt-tip pen. The participants were handed similar white A5 cards and the same pen to produce their drawings.

2.3. Procedure

Each participant was asked to examine a drawing for 10 s. After this time, the experimenter either removed the drawing from sight (in the Memory condition) or left it in sight (in the Copy condition) and handed the participant a blank card and a felt-tip pen as she asked him to copy the drawing as accurately as possible. The process of handing the card and paper took about 3–4 s in both conditions. Participants were not told what their task would be after looking at the drawing for 10 s in either condition. When he or she had finished drawing, the participant was asked to write down a brief description of the drawing he saw on another piece of paper (this was not analyzed in the current study).

We ran a total of 16 transmission chains, eight in the Memory and eight in the Copy condition, with 22 participants each. Half of the chains in each condition were initialized with one of the drawings in Fig. 1, and half with the other. The first participant in each chain was shown the initial drawing and the rest were shown the drawing produced by the previous participant in the chain.

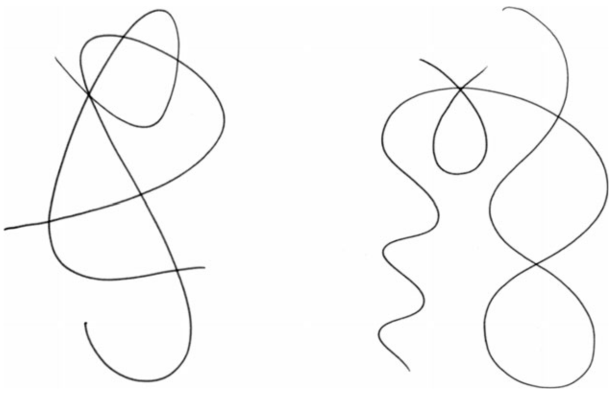


Fig. 1. The two initial drawings in the transmission chains.

2.4. Analyses

2.4.1. Perimetric complexity

We measured the perimetric complexity (PC) of each drawing following the algorithm in Pelli, Burns, Farell, and Moore-Page (2006), namely the squared length of the inside plus outside perimeters of a drawing divided by the ink area contained within those perimeters. Pelli et al. show that this complexity metric is an accurate negative correlate of the efficiency of object recognition. The complete process of each drawing involved scanning, saving in PNG format with maximum contrast (black on white only), and calculating PC.

Perimetric complexity is scale-invariant, provided the line thickness scales up with the size of the drawing. Scaled-up versions of a drawing will yield the same PC, but the same drawing produced in different sizes with the same pen (same line thickness) yields PC values proportional to drawing size.

Our drawings decreased in size over rounds, especially in the Memory condition (see Fig. 2a and b). Given that the drawings were all executed with the same pen, they had varying ratios of line thickness to overall drawing size. This compromises the accuracy of PC as a measure of complexity; therefore, we used an additional measure that proved scale-independent for our drawings, and therefore insensitive to the size confound—algorithmic complexity.

2.4.2. Algorithmic complexity

We estimated a proxy of the algorithmic or Kolmogorov complexity of each drawing by measuring the size in bytes of the compressed file of each drawing after several transformations, aimed at giving the smallest representation of the drawing that we could: First, they were scanned and saved in PNG format; the resulting drawings were then cropped to their bounding boxes, which were resized to low-resolution and identical areas (10,000 pixels) and then vectorized with the POTRACE algorithm (Selinger, 2003) into

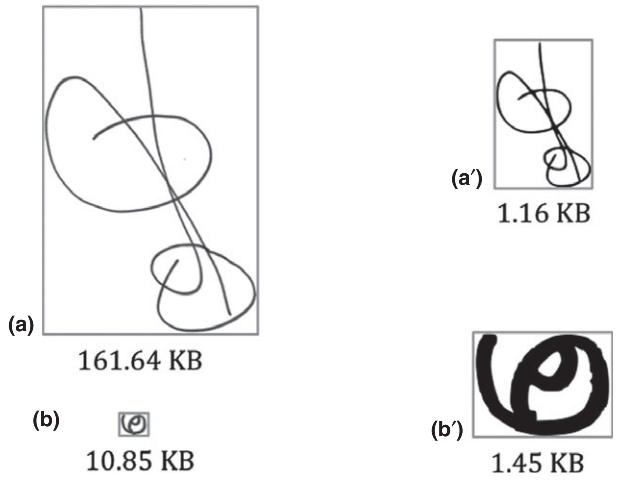


Fig. 2. Two of the drawings from the experiments in two different formats: left, cropped PNG format file; right, the resized vectorized version. File sizes in kilobytes given for each drawing.

FIG format. Finally, the vector representations (descriptions of the drawings in terms of shapes, curves, etc.) were zipped to get rid of any remaining simple redundancies. The result is an impressively compressed format, but the figures still look very close to the originals when displayed (see Fig. 2). The rescaling step sometimes introduced pixelation-style artifacts, which show up in the vectors. Specifically, some of the very small drawings are actually up-sized rather than down-sized For instance, the smaller drawing in Fig. 2(b) is enlarged at resizing and small ink imperfections in the edges of the line are blown up (Fig. 2b). In contrast, the larger drawing (Fig. 2a) is reduced at resizing and equivalent ink imperfections disappear. This artifact of the method means that the vectorized file descriptions of smaller drawings—but not those of larger ones—may end up longer than they should be. See, for instance, drawings (a) and (b) in Fig. 2: If we take the original PNG files, (a) is much larger than (b), but when we consider the vectorized files, then (b′) is actually a bit larger than (a′). This biases the results in favor of increased complexity in smaller drawings; in other words, our metric of complexity is likely to underestimate the difference between larger and smaller drawings, and, since drawings tended to be larger in the Copy condition than in the Memory condition, any difference between conditions will also be underestimated. As such, this is a conservative measure, given our hypothesis.

Complexity (the reverse of compressibility) was quantified as the size of the cropped vectorized files in bytes.

Perimetric and algorithmic complexity scores were submitted to repeated-measures ANOVAS with Reproduction condition (Memory or Copy) as between-subjects factor and Round (0–22) as within-subjects factor. Even though each drawing was produced by a different individual, it is appropriate to treat Round as within-subjects here because units of analysis in transmission chain studies are chains, not individual participants. Initial drawing (1 or 2) was included as an additional between-subjects factor to check for its effects. Cumulativeness over rounds was calculated with Page's trend test; overall change in complexity was calculated comparing the values at the initial and final rounds with a paired t test.

2.4.3. *Conventionality*

Six people rated all the drawings each for whether they represented something recognizable or not on a 7-point Likert scale. Judging was done in an online form where all 354 drawings (22 in each of 16 chains plus the two initial drawings in Fig. 1) appeared in sequence, in a random order, on the screen next to a 7-point Likert scale. The instructions specified that 1 should be given if they thought the drawing represented something conventionally recognizable (including symbols like letters and numbers), in other words, if they were confident that different people would agree on what the drawing represented. On the opposite end of the scale, 7 should be given if they thought it did not represent anything conventionally recognizable and they thought different people might give divergent interpretations of what the drawing represented. With this question, we tried to avoid people rating the drawing as recognizable if they happened to see something in the drawing that was meaningful to them in an idiosyncratic way—almost any drawing can be creatively interpreted as something recognizable, but we were trying to measure *conventional* recognizability.

The inter-rater agreement was calculated, and Mann–Whitney and Wilcoxon rank tests were performed to examine the effect of Reproduction condition (Memory vs. Copy) as well as Initial drawing (1 vs. 2) on conventionality rates.

3. Results

Inspection of the drawing chains (Fig. 3) impressionistically reveals that drawings in both conditions do change, but in different ways. In the Memory condition, they become less complex, tend to turn into conventional symbols like letters and numbers, and become smaller (see Fig. 4, left) over rounds of transmission. In the Copy condition, however, drawings maintain the original complexity to a higher degree and remain meaningless.

3.1. *Complexity*

The ANOVA on the PC values revealed significant main effects of Round ($F(22, 264) = 16.030$; $p < .001$), Reproduction condition ($F(1, 12) = 51.788$; $p < .001$) and initial drawing ($F(1, 12) = 18.292$; $p < .01$), and significant Round × Reproduction condition ($F(22, 264) = 2.405$; $p < .01$) and Round × Initial drawing ($F(22, 264) = 4.829$; $p < .001$) interactions.

The main effect of Reproduction condition is due to less complex drawings being produced in the Memory ($M = 731.4$, $SD = 852.2$) than in the Copy conditions ($M = 2{,}069$

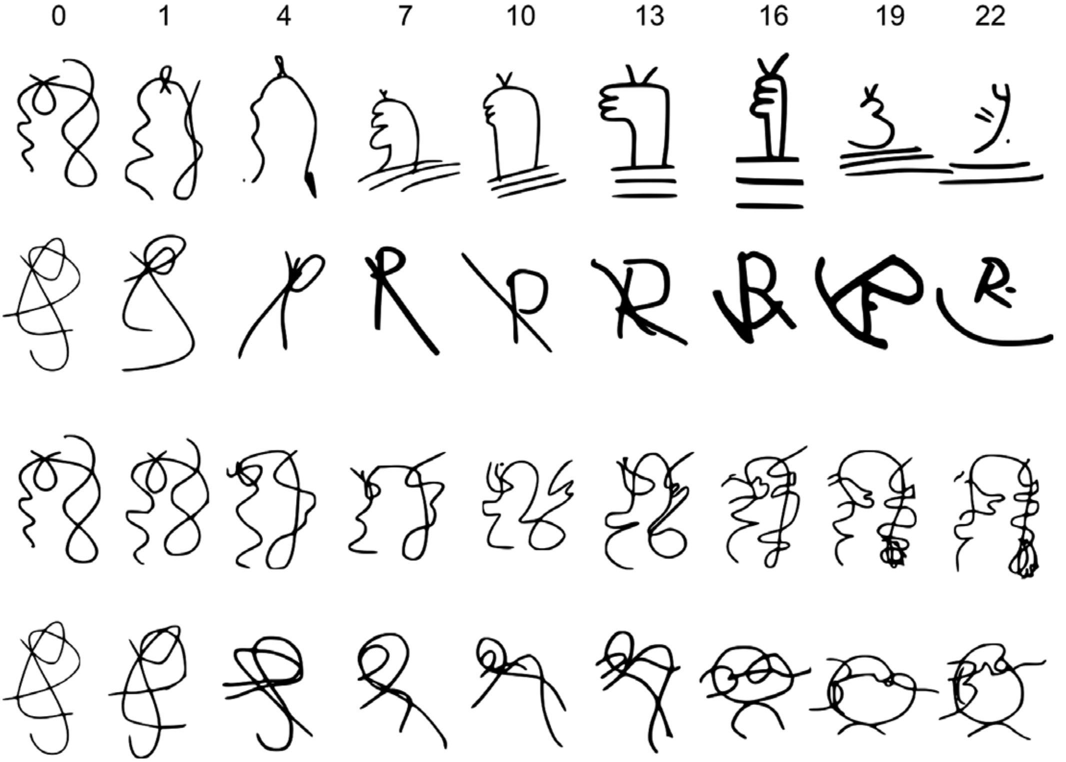


Fig. 3. A selection of resized, vectorized drawings from the experiments, in two of the Memory chains (top) and two of the Copy chains (bottom). Initial drawings (0) plus drawings at rounds 1, 4, 7, 10, 13, 16, 19, and 22 are shown. The different thickness of the lines is due to the original drawings being resized: smaller drawings are enlarged, and therefore have thicker lines.

SD = 1,083; see Fig. 4, right). The main effect of Round is due to an overall decrease in PC over rounds. This decrease is cumulative, as indicated by significant Page's *L* ($m = 16$, $n = 23$, $L = 62{,}087$, $p < .001$) and *t* test ($t(15) = 4.322$, $p < .001$).

The effect of Initial drawing is due to chains initialized with drawing 1 (Fig. 1, left; Drawing 1: $M = 1{,}797.29$; $SD = 1{,}252.01$) achieving lower complexity values than chains initialized with drawing 2 (Fig. 1, right; Drawing 2: $M = 1{,}002.62$; $SD = 955.03$).

The Round × Reproduction condition interaction is illustrated in Fig. 4 (right). In order to explore this interaction further, separate repeated-measures ANOVA with Round as within-subjects factor were carried out in the Memory and Copy conditions. There was an effect of Round in the Memory condition ($F(22, 132) = 33.053$, $p < .001$), and an also significant, but much smaller one in the Copy condition ($F(22, 132) = 2.398$, $p < .01$). This difference is confirmed by *t* tests: The difference in complexity between initial and final rounds is significant in the Memory ($t(7) = 6.557$, $p < .001$), but not in the Copy condition ($t(7) = 1.643$, $p = .140$). These results indicate that the drawings in the Memory condition lose complexity faster than in the Copy condition.

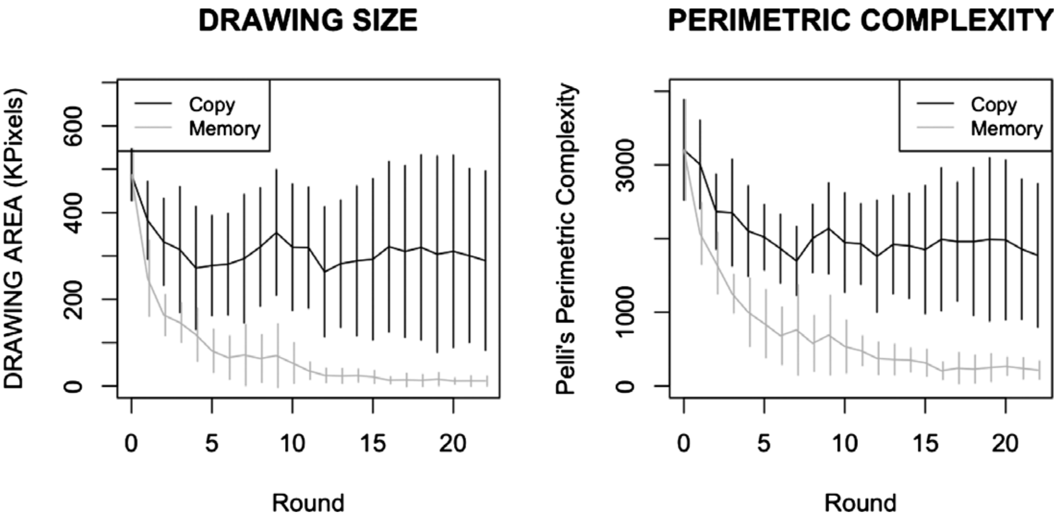


Fig. 4. Size (left) and perimetric complexity (right) of the drawings in the Copy and Memory conditions, by Round (95% confidence intervals shown).

As for algorithmic complexity values, the ANOVA revealed significant main effects of Reproduction condition ($F(1, 12) = 13.968$; $p < .01$) and Round ($F(22, 264) = 4.186$; $p < .001$), and a Round × Reproduction condition interaction ($F(22, 264) = 2.662$; $p < .01$). There was no main effect of Initial drawing, nor any interactions of this factor with any others, further indicating that, in our data, PC may be influenced by size while algorithmic complexity is not.

The main effect of Reproduction condition is due to less complex drawings being produced in the Memory ($M = 985.6$, $SD = 270.5$) than in the Copy conditions ($M = 1{,}356$, $SD = 286.5$). The main effect of Round is due to an overall decrease in file size over rounds. This decrease is cumulative, as indicated by a significant Page's L trend test ($m = 16$, $n = 23$, $L = 59{,}644$, $p < .01$), and there is a significant difference between values at rounds 0 and 22 ($t(15) = 2.862$, $p = .012$).

The interaction between Round and Reproduction condition is illustrated in Fig. 5. In order to explore this interaction further, separate repeated-measures ANOVA with Round as within-subjects factor were carried out in the Memory and Copy conditions. There was an effect of Round in the Memory condition ($F(22, 132) = 7.340$, $p < .001$), but not in the Copy condition ($F(22, 132) = 0.606$, $p = .866$). This difference is confirmed by t tests: In the Memory condition the decrease in complexity is significant ($t(7) = 12.26$, $p < .001$) and also cumulative: $m = 8$, $n = 23$, $L = 26{,}614$, $p < .001$. In the Copy condition, although the values descend cumulatively ($m = 8$, $n = 23$, $L = 28{,}848$, $p = .003$), the complexity at the final round is not lower than that at the starting round ($t(3) = 0.176$, $p = .866$). This analysis indicates that the interaction is due to an overall decrease in complexity over rounds in the Memory but not in the Copy condition.

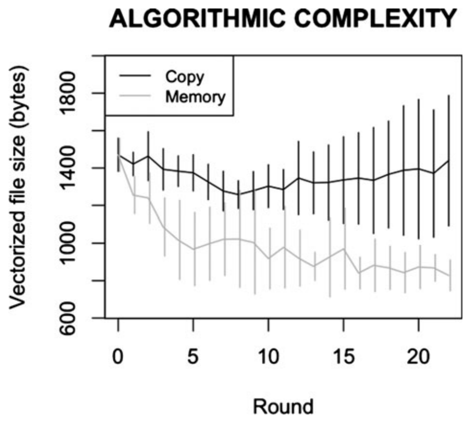


Fig. 5. Algorithmic complexity of the drawings in the Copy and Memory conditions, by Round (95% confidence intervals shown).

3.2. Conventionality

The Likert scale scores from the six raters showed a high inter-rater reliability (Pearson's R range (0.30–0.68), all $p < .001$).

Mann–Whitney and Wilcoxon rank tests on all scores collapsing across rounds revealed a significant effect of Reproduction condition on Likert scores ($U = 311{,}675$; $W = 866{,}606$ ($N = 1{,}047$), $p < .001$), with drawings in the Copy condition rated as more abstract than those in the Memory condition (Fig. 6, left). There was also a significant effect of Initial drawing ($U = 476{,}979.5$, $W = 1{,}032{,}964.5$, $p < .001$), with drawings from chains initialized with drawing 2 were rated as more abstract than drawings in chains initialized with drawing 1.

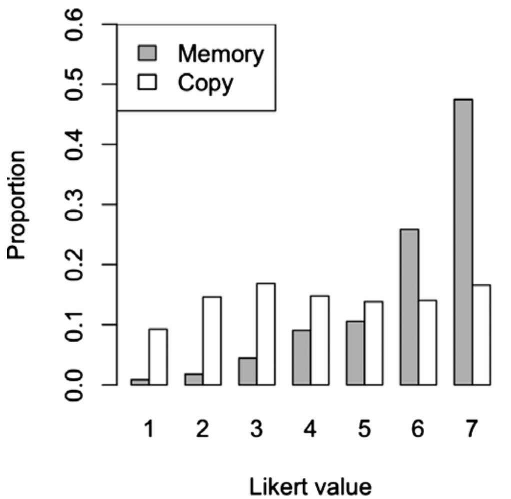
Mann–Whitney and Wilcoxon tests were performed to examine the effect on Likert scores for each of the 22 rounds of Reproduction condition (Table 1; see Fig. 6, right for median values) and Initial drawing (Table 2).

The results in Table 1 indicate that the rates accorded to the drawings in the Memory and the Copy conditions are similar until round 2, after which they become significantly different. The initially arbitrary drawings began to be rated as significantly more conventional in the Memory condition at round 3, while in the Copy condition they continued to be rated as highly arbitrary until the final round.

The results in Table 2 indicate that the rates accorded to the drawings in chains initialized with both drawings are similar in all rounds except 10 and 17–22.

4. Discussion

The algorithmic complexity and conventionality results indicate that the manipulation in the Memory condition—the learning aspect of cultural transmission, namely learning a



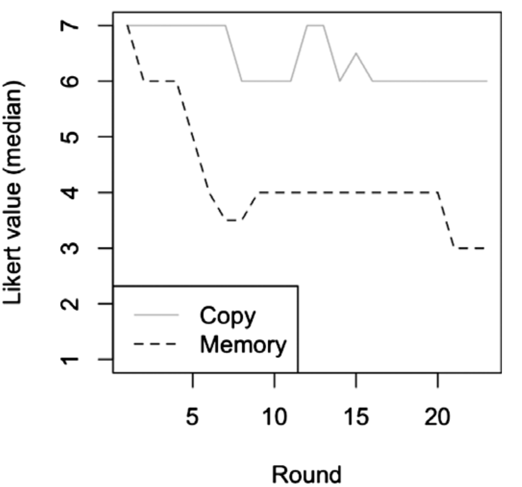


Fig. 6. Left: distributions of Likert scale values 1–7 given to the drawings in the Memory and Copy conditions, collapsing across Rounds. Right: median Likert scale scores given to drawings in rounds 0–22, by Reproduction condition. Higher values indicate the drawings are rated as highly arbitrary and lower values, as highly conventional.

Table 1
Mann–Whitney's U and Wilcoxon's W values and significance of reproduction condition (Copy or Memory) by round, calculated with 48 judgments per round per condition

Round	U	W	p	Round	U	W	p
1	879	2,055	=.045	12	521	1,679.5	<.001
2	890	1,971	=.115	13	735	1,863	<.003
3	722	1,850.5	=.002	14	746	1,922	=.002
4	598	1,774.5	<.001	15	588	1,716	<.001
5	471	1,647	<.001	16	636	1,812.5	<.001
6	585	1,709	<.001	17	578	1,659	<.001
7	629	1,805	<.001	18	504	1,494	<.001
8	584	1,712.5	<.001	19	622	1,798.5	<.001
9	726	1,902.5	<.002	20	448	1,624	<.001
10	649	1,825	<.001	21	519	1,600.5	<.001
11	579	1,755.5	=.000	22	506	1,682	<.001

drawing and holding it in memory for a few seconds—resulted in an increase in compressibility (low complexity and high conventionality). The Copy condition, where this small amount of learning was absent because the original drawing was in full view during reproduction, yielded very different results: The items retained the complexity and arbitrariness of the random drawings used to initialize the chains.

The increase in compressibility in the Memory condition came about in at least two ways. First, by losing complexity: The drawings were gradually simplified in form by

Table 2
Mann–Whitney's U and Wilcoxon's W values and significance of Initial drawing (1 or 2) by round, calculated with 48 judgments per round per condition

Round	U	W	p	Round	U	W	p
1	1,034	2,210	=.359	12	1,092.5	2,268.5	=.651
2	982.5	2,110.5	=.247	13	991	2,167	=.230
3	1,082	2,136	=.861	14	917.5	2,093.5	=.108
4	1,120	2,296	=.809	15	912	1,993	=.139
5	1,024	2,200	=.334	16	991.5	2,119.5	=.300
6	1,110	2,286	=.756	17	769.5	1,945.5	=.006
7	1,052	2,248	=.550	18	792.5	1,873.5	=.035
8	1,094.5	2,270.5	=.799	19	821.5	1,997.5	=.013
9	932	2,108	=.136	20	737	1,913	=.002
10	810.5	1,986.5	=.016	21	662.5	1,838.5	<.001
11	973	2,149	=.117	22	718	1,894	=.001

becoming smaller, and by being made up of fewer and shorter lines (as in Fay, Garrod, Roberts, & Swoboda, 2010; Garrod et al., 2007) and thus became more learnable. This is what experiments and computer models carried out in the last few years predict (Kirby, 2001; Kirby et al., 2007, 2008; Smith et al., 2013; Theisen-White et al., 2011; Verhoef et al., 2011). Second, transforming into conventional signals like letters or numbers (as cognitive theories of memory would predict, e.g., Bartlett, 1932; Fillmore, 1977; Langacker, 1987; Minsky, 1975; Schank & Abelson, 1977) may also contribute to compressibility. Our approximation of Kolmogorov complexity is based on the length of description of the lines in the drawings, but it does not take into account any meaning the drawings may convey. However, if we had used an algorithm that was able to understand about letters, numbers and other recognizable patterns, the description of some of the drawings could be as short as "letter R with a line and a dot underneath" (Fig. 3, second chain), in comparison with the much longer descriptions that would be required for the last drawings in the two copy condition chains at the bottom of Fig. 3.

The drawings in the Copy condition also changed, but for different reasons than those in the Memory condition. The innovations introduced here were random, in the sense that they were unbiased with respect to pre-existing meanings and with respect to complexity. They may have been caused by production error—we would in fact predict that if the participants in this condition were trained artists, the drawings would change at a slower pace because there would be fewer production errors. Some of these random innovations were reproduced in an unbiased way by the following generations. The effects of this reproduction modality are introducing and retaining random mutation as happens in neutral evolutionary dynamics, or drift.

In the Memory condition, on the contrary, not all innovations had an equal chance to be reproduced: Those that were better adapted to the compressibility biases would be preferentially reproduced, or selected for, and the least fit in this sense would be selected against. Once the cultural system attained an optimal level of simplicity and conventionality, the pressures derived from keeping something in memory ceased to act, and the process of adaptation ended. Figs. 5 and 6 illustrate this limited selection effect, where complexity and arbitrariness values, respectively, undergo a rapid descent before stabilizing in the Memory condition. This is also what happens in studies that initialize transmission chains with a random language (Kirby et al., 2008; Perfors & Navarro, 2014), or a random set of items (Kalish, Griffiths, & Lewandowsky, 2007; Verhoef et al., 2011), which cumulatively evolve to become highly compressible, and then stabilize.

This study thus shows that culturally transmitted information is affected differently by different aspects of social transmission. First, it provides an original demonstration that direct *reproduction* can introduce innovations that are random with respect to cognitive biases, and retain those innovations across the generations. This random-mutation-and-retention process can maintain complexity and arbitrariness in cultural traditions. Second, the study crisply demonstrates that when behavior is transmitted through the bottleneck of *learning* (even if this bottleneck is as tiny as remembering a drawing for a matter of seconds), then we see a cumulative increase in compressibility.

Acknowledgments

The authors are thankful to Ana Paula Arribas Bermejo for her help with data collection and acknowledge support from AHRC grant number AH/F017677/1.

References

Bangerter, A. (2000). Transformation between scientific and social representations of conception: The method of serial reproduction. *British Journal of Social Psychology*, *39*, 521–535.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Brighton, H., Kirby, S., & Smith, K. (2005). Language as an evolutionary system. *Physics of Life Reviews*, *2*, 177–226.

Chater, N., & Vitanyi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences*, *7*(1), 19–22.

Esper, E. A. (1925). *A technique for the experimental investigation of associative interference in artificial linguistic material*. Philadelphia, PA: Linguistic Society of America.

Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communicative systems. *Cognitive Science*, *34*, 351–386.

Fillmore, C. (1977). Topics in lexical semantics. In R. W. Cole (Ed.), *Current issues in linguistic theory* (pp. 76–138). Bloomington, IN: Indiana University.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767.

Garrod, S., Fay, N., Leed, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, *31*, 961–987.

Heyes, C. M. (2012). Grist and mills: On the cultural origins of cultural learning. *Philosophical Transactions of the Royal Society B*, *367*, 2181–2191.

Heyes, C. M., Huber, L., Gergely, G., & Brass, M. (Eds.) (2009). Evolution, development and intentional control of imitation (whole issue). *Philosophical Transactions of the Royal Society B*, *364*, 2291–2443.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin Review*, *14*, 288–294.

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin*, *26*, 594–604.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, *5*, 102–110.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, *104*(12), 5241–5245.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, *97*(3), 405–423.

Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision*, (pp. 211–277). New York: McGraw-Hill.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646–4674.

Perfors, A., & Navarro, D. J. (2014). Language evolution can be shaped by the structure of the world. *Cognitive Science*, *38*, 775–793.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale NJ: Erlbaum.

Selinger, P. (2003). *POTRACE: A polygon-based tracing algorithm (software).*

Smith, K., Tamariz, M., & Kirby, S. (2013). Linguistic structure is an evolutionary trade-off between simplicity and expressivity. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1348–1353). Austin, TX: Cognitive Science Society.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford, UK: Blackwell.

Theisen-White, C., Kirby, S., & Oberlander, J. (2011). Integrating the horizontal and vertical cultural transmission of novel communication systems. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 956–961). Austin, TX: Cognitive Science.

Verhoef, T., Kirby, S., & Padden, C. (2011). Cultural emergence of combinatorial structure in an artificial whistled language. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 483–488). Austin, TX: Cognitive Science.